The loop E–loop D region of *Escherichia coli* 5S rRNA: the solution structure reveals an unusual loop that may be important for binding ribosomal proteins

Anne Dallas[1] and Peter B Moore[1, 2*]

Background: 5S ribosomal RNA is the smallest rRNA. Its Watson–Crick helices were identified more than 20 years ago, but the conformations of its loops have long defied analysis. One of the three arms of 5S rRNA, residues 69–106 in *Escherichia coli*, contains a 14-residue internal loop called loop E. The sequence of loop E is conserved within kingdoms, and is terminated by a pyrimidine-rich loop called loop D. Loop E is the binding site for the ribosomal protein L25 in the *E. coli* ribosome.

Results: The solution structure of a 42-nucleotide derivative of *E. coli* 5S rRNA that includes loops D and E has been determined by nuclear magnetic resonance spectroscopy. Formally, loop E is not a loop at all; it is a double helical structure that contains seven, consecutive non-Watson–Crick base pairs. The major groove of the molecule is narrowed in loop E, and an unusual array of hydrogen-bond donors and acceptors appear in its minor groove. Loop D, which on paper looks like a three-pyrimidine terminal loop closed by a GC, is better thought of as a five-base loop because its closing GC is not a normal Watson–Crick pair. The two pyrimidines on the 5′-side of the loop are stacked on each other, and tilt into the minor groove of the adjacent helix. The third pyrimidine is fully exposed to solvent.

Conclusions: This structure rationalizes all the biochemical and chemical protection data available for the loop E–loop D arm of intact 5S rRNA. While the molecule is double helical over its entire length, the geometry of its internal loop is highly irregular, and its irregularities may explain why the loop E–loop D arm of 5S rRNA interacts specifically with ribosomal protein L25 in *E. coli*.

Addresses: [1]Department of Chemistry, Yale University, PO Box 8107, New Haven, CT 06520-8107, USA and [2]Department of Molecular Biophysics and Biochemistry, Yale University, PO Box 8107, New Haven, CT 06520-8107, USA.

*Corresponding author.
E-mail: moore@neutron.chem.yale.edu

Key words: loop E, NMR, ribosome, 5S rRNA, structure

Received: **10 September 1997**
Revisions requested: **8 October 1997**
Revisions received: **23 October 1997**
Accepted: **24 October 1997**

Structure 15 December 1997, **5**:1639–1653
http://biomednet.com/elecref/0969212600501639

Introduction

5S rRNA is an ~120 nucleotide RNA found in virtually all ribosomes. It is located in the central protuberance of the large ribosomal subunit near the peptidyl transferase and factor-binding sites, but its function is unknown [1–3]. Because 5S rRNA is small and easy to prepare, over the years, RNA chemists have devoted an effort to its characterization disproportionate to its size. It has been established that the secondary structure of 5S rRNA consists of three largely helical arms radiating out from a common junction (Figure 1a) [4]; little, however, is known about the architecture of its loops.

The conformation of the arm of *Escherichia coli* 5S rRNA that contains loops D and E has been under investigation in this laboratory by NMR spectroscopy for over a decade. Loop D is the three-pyrimidine loop that terminates the helix IV–V arm of 5S rRNA. Its residues can be cross-linked to 23S rRNA near the peptidyl transferase center [5,6]. Loop E is a 14-residue internal loop that is part of the binding site of ribosomal protein L25 in *E. coli* [7].

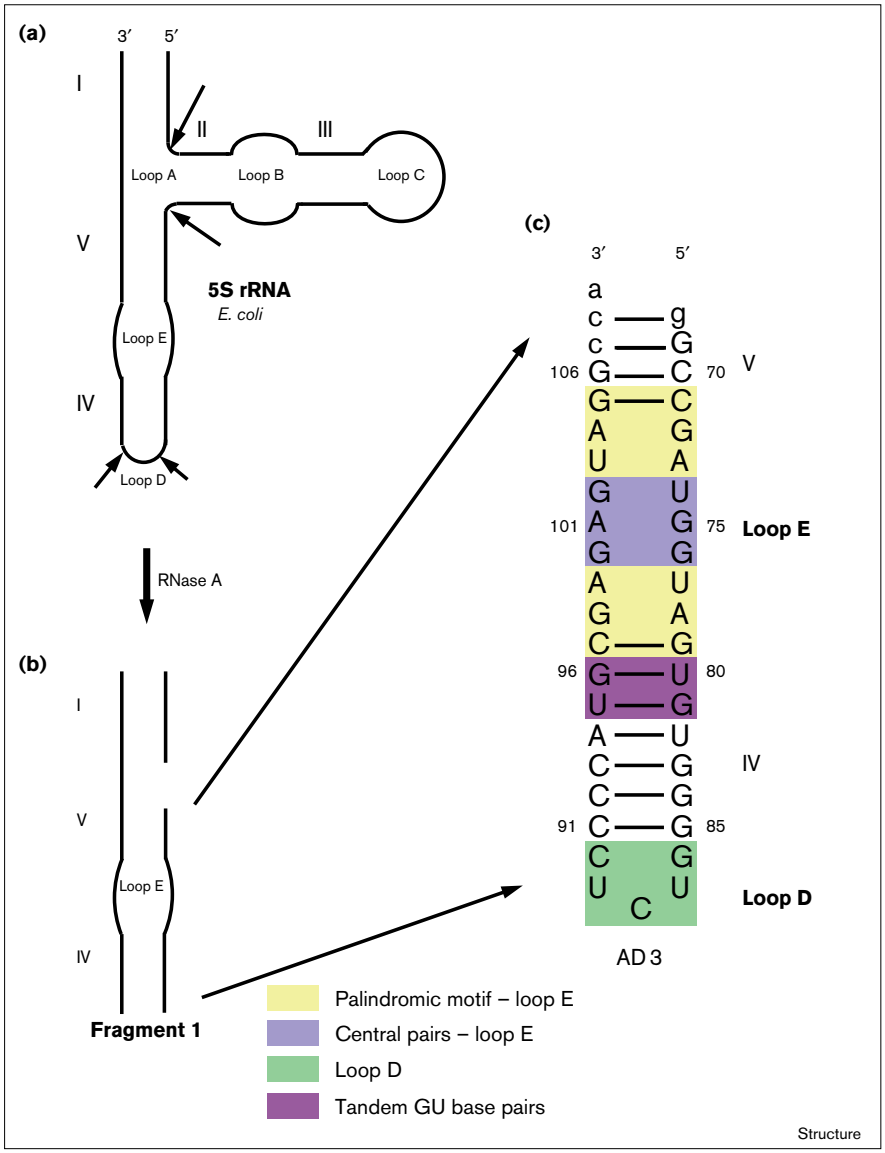
Initial NMR studies of a 62-base nucleolytic fragment of 5S rRNA called fragment 1 [8] (Figure 1b) demonstrated that loop E is highly structured in solution, but it was not possible to assign the spectra of fragment 1 in sufficient detail to derive a model for it. In order to obtain further information, a truncated version of fragment 1, called AD3, was produced that lacks helix I, the structure of which had already been determined [9]. Preliminary spectroscopic studies demonstrated that this 42-base RNA, which contains loop E, loop D, and helices IV and V (Figure 1c), has a conformation similar to that of the same sequence in fragment 1, and like intact 5S rRNA binds ribosomal protein L25. A proposal for the conformation of loop E emerged in 1995 [10]. Here we present the solution structure of the entire AD3 molecule, including loop D. The conformational peculiarities of the loop E–helix IV region may explain why 5S rRNA is able to make specific interactions with other macromolecules.

Results

Ionic conditions

The imino proton spectrum of AD3 is a subset of the spectrum of 5S rRNA and of the spectrum of fragment 1

Figure 1



Schematic diagram of the 5S rRNA, fragment 1 and the sequence of the 5S rRNA derivative used in this study, AD3. **(a)** Secondary structure of *E. coli* 5S rRNA. Arrows indicate sites where RNase A cleaves under mild digestion conditions. Roman numerals designate helical regions and letters identify loops. **(b)** Fragment 1. **(c)** The sequence of AD3. Lowercase letters denote residues added to G69–G106 of *E. coli* 5S rRNA to facilitate transcription by T7 RNA polymerase. Noncanonical sequence features are highlighted in color (see key at the bottom of the figure).

[10–12], as expected, but because AD3 is smaller than both of these molecules its spectra contain fewer resonances and their line widths are narrower. It was known from previous work on fragment 1 that divalent cations stabilize the conformation of loop E [13], and that the imino proton resonances related to loop E are better resolved in buffers that contain Ca^{2+} than they are in buffers that contain Mg^{2+} [12]. Similar studies done with AD3 confirmed these results. In the absence of divalent cations, all the imino proton resonances associated with loop E either broaden dramatically or disappear entirely (data not shown). In an effort to further optimize the imino proton spectrum of AD3, its appearance was studied as a function of pH. As Figure 2 demonstrates, reduction of the pH from 7.5 to 5.5 narrows resonance line widths, and improves resolution in the critical 10–11 ppm region of the AD3 spectrum.

These findings notwithstanding, there is no reason to believe that either the replacement of Mg^{2+} with Ca^{2+} or the reduction of the pH from 7.2 to 5.5 significantly alters the structure of AD3. For example, when the pH drops from 7.2 to 5.5, the chemical shifts of non-exchangeable proton resonances of AD3 change by 0.02 ppm or less, which is the accuracy with which chemical shifts can be measured in the first place. The effect of replacing Mg^{2+} with Ca^{2+} is similarly small. In addition, under all conditions tested the double quantum-filtered correlated spectroscopy (DQF-COSY) spectrum of AD3 contains a single H5/H6 cross-peak for each of its 18 pyrimidines, which

suggests the molecule has a unique conformation over the entire range of conditions explored (data not shown). Most of the spectra discussed below were taken in Ca^{2+} at acid pH, the conditions that optimize the molecule's imino proton spectrum.

Assignment of imino proton resonances

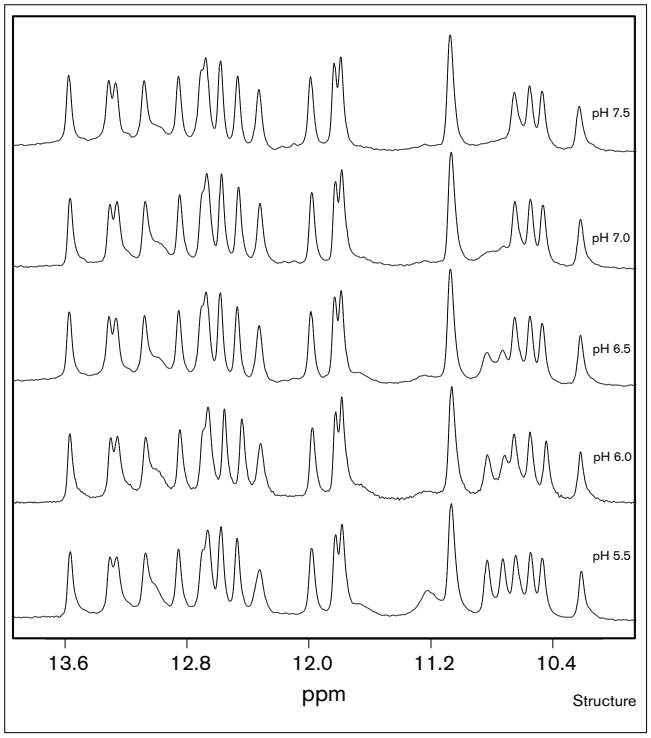
A ^{15}N-^{1}H heteronuclear multiple quantum coherence (HMQC) spectrum of a fully labeled $^{15}N/^{13}C$ sample distinguished the GH1 from the UH3 imino proton resonances in the downfield spectrum of AD3 [14]; there are 15 of the former, and six of the latter (Figure 3a). The imino proton resonances originating from helix V, helix IV and the two GUs adjoining helix IV were assigned using correlations observed in H_2O nuclear Overhauser effect spectroscopy (NOESY) spectra (Figure 3b). The assignments obtained correspond to those reported earlier for the corresponding segments of fragment 1 [12,15].

Two years ago, we published a partial set of assignments for the imino proton resonances that were based on data obtained with unlabeled samples, before the effects of ionic conditions and pH on the downfield spectrum of AD3 had been fully explored [10]. Data obtained at acid pH have forced us to reassign three of these resonances, all of which originate in loop E, and have also enabled us to complete the assignment of the AD3 downfield spectrum.

There are five Gs and three Us in the loop E sequence that could contribute to the downfield spectrum of AD3. The assignments of five of them were unaffected by the new data: U74, G76, U77, G98 and U103 [10]. The assignments of the UH3 resonances in question were easily confirmed because they give nuclear Overhauser effects (NOEs) to their own H5 and H6 protons, which could be assigned independently (see below). The G98 H1 resonance was assigned on the basis of its NOEs to the imino protons of U77 and G79. The assignment of G76s imino proton was inferred from NOEs to the imino and anomeric protons of U77 as well as to the H2 of A101.

This left the GH1 resonances that appear at 10.5 ppm, 10.6 ppm and 10.8 ppm to consider, which were originally assigned to G72, G102 and G75, respectively. The resonance at 10.5 ppm gives an NOE to U103 H3, which would be plausible for either G72 H1 or G102 H1 (Figure 3b), but a H_2O-NOESY spectrum collected at 5°C revealed an additional weak NOE that correlates it to U74 (see Figure 1c). Thus the resonance at 10.5 ppm must belong to G102, the residue opposite U74, and not G72. The resonance at 10.6 ppm gives weak NOEs to both the hydrogen bonded and non-hydrogen bonded amino protons of C71 and an NOE to A104 H8, which indicates that it belongs to G72 and not G102 (data not shown). The 10.8 ppm resonance is more problematic. This resonance gives weak NOEs to G76 H1, A101 H2 and A101 H1′, and could therefore represent the imino proton of either G100 or G75. As the rest of the spectra indicate that loop E is helical throughout its length (see below), we favor the assignment of the 10.8 ppm resonance to G100, which leaves the imino proton of G75 unaccounted for. This proton must exchange too rapidly with the solvent to be observed.

Figure 2

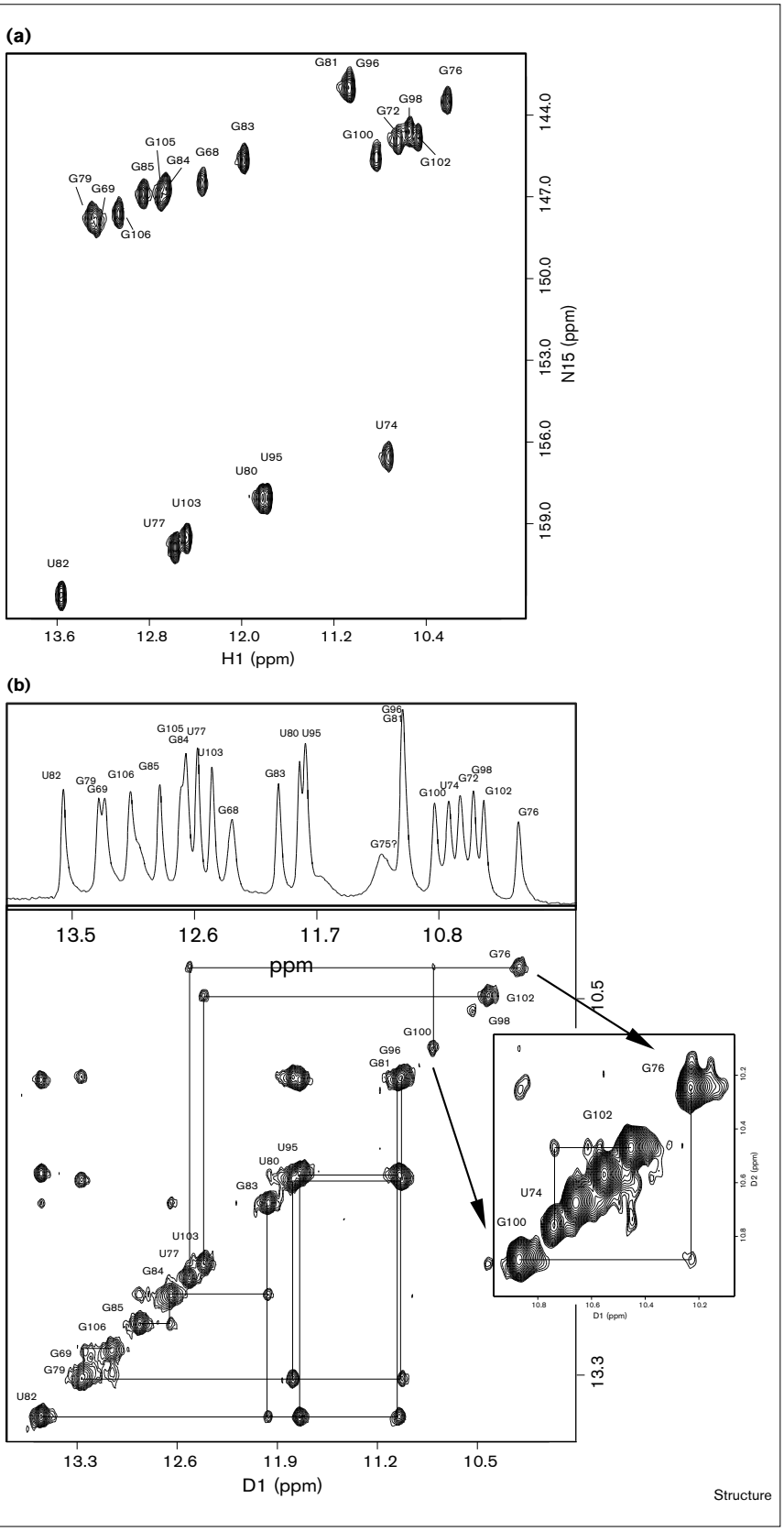


The imino proton spectrum of AD3 as a function of pH at 10°C. The pH of the sample used for each spectrum was set by dialysis against an appropriate buffer.

Assignment of non-exchangeable protons

Non-exchangeable proton spectra of AD3 were acquired both in buffers containing Mg^{2+} and in buffers containing Ca^{2+}, and assignments were made at both 30°C and 10°C. The anomeric–aromatic region of AD3's $^{2}H_2O$-NOESY spectrum contains an H1′/(H6 or 8) sequential walk [16] that runs from its 5′-end to its 3′-end with a single break between C88 H6 and U89 H1′, at the apex of loop D. In Figure 4a, the 5′-side of the walk is shown in blue, and its 3′-side is shown in red. The only features of the AD3 walk that distinguish it qualitatively from that of an A-form helix are the chemical shifts of the anomeric protons of G79 and G105, which belong to the GC base pairs that close loop E. They resonate about 1 ppm upfield of normal. The identities of these resonances were confirmed using a ^{1}H-^{13}C constant time heteronuclear single quantum coherence (CT-HSQC) experiment

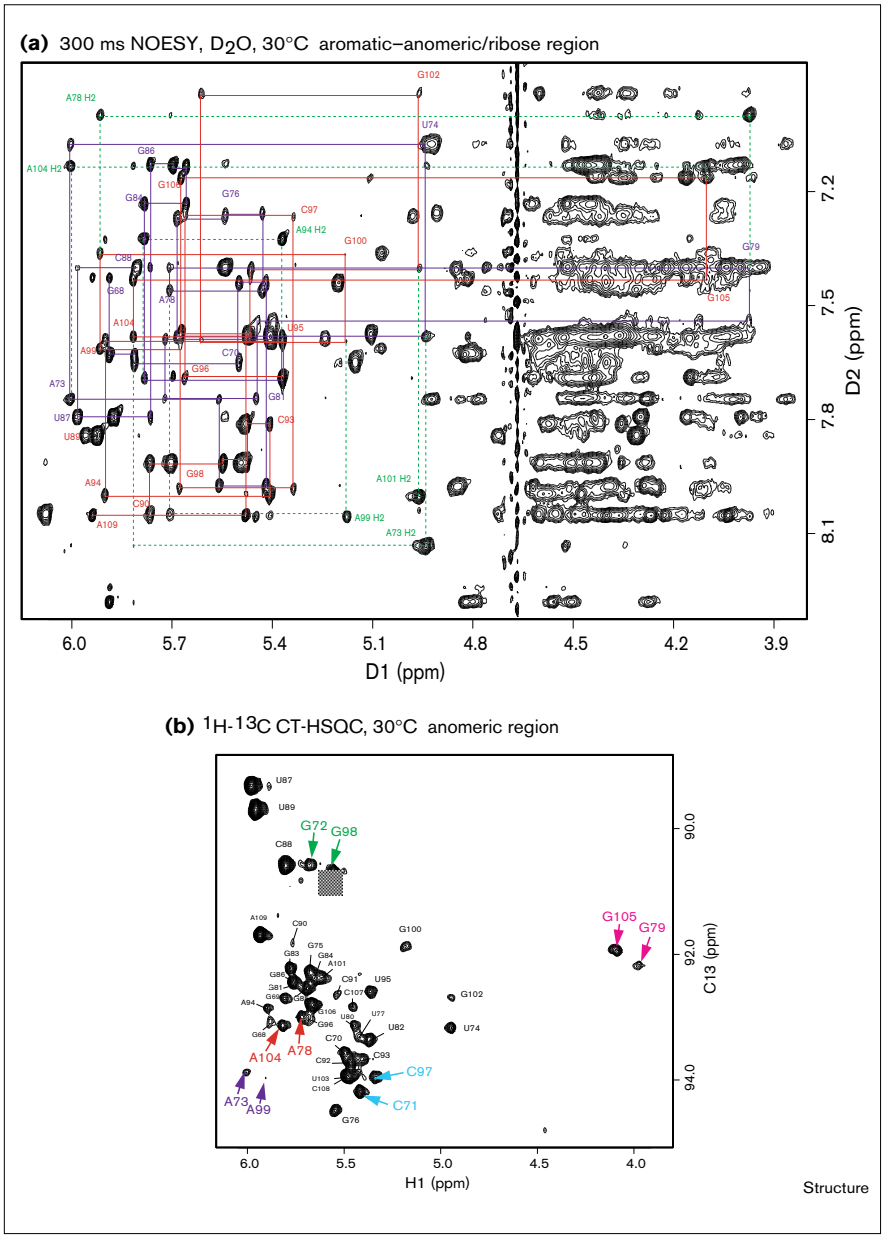
Figure 3



Exchangeable proton spectra of AD3. **(a)** Imino resonance region of a ^{15}N-^{1}H HMQC at 10°C. GH1 and UH3 ^{15}N-^{1}H cross-peaks are labeled with their assignments. **(b)** The imino region of a 150 ms H_2O-NOESY spectrum collected at 10°C. The trace of the 1D imino spectrum is included along the top of the 2D spectrum. Autopeaks are labeled along the diagonal. Lines connect pairs of protons that are correlated through space. The inset shows the region between 10–11 ppm collected at 5°C. An additional cross-peak is observed under these conditions between G102 H1 and U74 H3 protons.

Figure 4

Spectra of non-exchangeable aromatic and anomeric resonances of AD3. **(a)** The aromatic–anomeric/ribose region of a 300 ms NOESY collected at 30°C. The purple line denotes the aromatic–anomeric walk from residues 68 to 88 (the 5′-side of the molecule). The red line traces the aromatic–anomeric connectivities up the 3′-side of AD3 from residues 89 to 109. The dashed green lines correlate NOEs involving adenine H2 protons. **(b)** The anomeric region of a ^{1}H-^{13}C CT-HSQC showing one-bond correlations between H1′-C1′ nuclei in AD3. The chemical shifts of the two upfield anomeric protons, G79 and G105 have been labeled in magenta. Several other pairs of anomeric cross-peaks of symmetry-related residues in loop E are also highlighted.



run on a ^{13}C/^{15}N-labeled sample of AD3 (Figure 4b) [17]. It should be noted that we were unable to assign the cross-peaks at 8.25 ppm definitely. These resonances were not reproducible from one sample to the next, and were not observed in this particular sample in the spectra acquired immediately after it was synthesized. They are likely to represent an alternate conformer of the dangling 3′-terminal residue, A109.

Adenine H2s were identified using one-bond C2-H2 correlations, which were observed in a ^{1}H-^{13}C CT-HSQC spectrum, and by identifying two-bond N1-H2 and N3-H2 correlations in a ^{15}N-^{1}H heteronuclear multiple bond correlation (HMBC) experiment [18]. Although the signal-to-noise was poor because of AD3s short transverse relaxation times (T_2s), it was possible to identify six of the seven expected H2-H8 correlations using an HccH total correlation spectroscopy (TOCSY) experiment [19]; the seventh was assigned using NOESY connectivities [16].

H2′ resonances were assigned using an HccH COSY [20] experiment that correlated H1′ and H2′ resonances from the same ribose ring. These assignments were confirmed using a CT-hCcH COSY [21] that displayed C1′ and H2′

cross-peaks for each ribose. Complete assignment of H3′, H4′, H5′ and H5″ resonances could not be achieved by heteronuclear methods because of the poor dispersion of C3′, C4′ and C5′ resonances in the AD3 ^{13}C spectrum. Even a three-dimensional HcCH TOCSY [22] spectrum did not resolve ribose spin systems unambiguously, except for those in loop D. As pointed out earlier [10], T_2s are longer in loop D, presumably because of local flexibility, and resonances originating in that part of the molecule tend to be in outlying regions of the spectra. Thus our ribose assignments beyond H2′ rely heavily on NOESY data and are incomplete.

Assignment of ^{31}P resonances

It was reported previously that the ^{31}P spectrum of fragment 1 is not well resolved, and that the few downfield-shifted resonances it contains belong to loop D [23]. It was no surprise, therefore, to find that all the ^{31}P nuclei in AD3, except two, resonate in the region characteristic of A-form RNA helices (data not shown) [24,25]. The two outlying resonances were assigned using H3′-P couplings detected in a ^{1}H-^{31}P COSY experiment [26] and belong to loop D: G86-P-U87 and C88-P-U89.

Direct conformational inferences

A great deal can be learned about the conformation of AD3 by inspection of its spectra. First, there are no intra-nucleotide aromatic–anomeric cross-peaks in its NOESY spectra comparable in intensity to a pyrimidine H5/H6 cross-peak. This means that the glycosidic torsion angles (χ) of all AD3 residues lie in the *anti* range [27]. Second, there are only a few cross-peaks observed in the H1′-H2′ region of its DQF-COSY spectrum (data not shown). The weakest of them belongs to the unpaired 3′-terminal residue, a109, which indicates that it samples different ribose conformations. The peaks that have magnitudes that imply the riboses are in the *2′-endo* conformation belong to loop D residues U87, C88 and U89. All residues in helix V, helix IV and loop E have ribose rings with C3′-*endo* pucker as the $^{3}J_{H1'-H2'}$ are too small to be observed [28].

Furthermore, it is clear from the AD3 imino proton spectra that residues g68–C71 form a Watson–Crick paired double helix with G105–c108 (helix V), and that G79–G85 is similarly paired with C91–C97. The two GUs contained in the latter helix are wobble base pairs (helix IV). In addition, as pointed out before, the (AG)(UA) pairs at the two ends of loop E are a type II, sheared AG and a reversed Hoogsteen UA, respectively [10]. Similar motifs have been studied in the eukaryotic loop E [29] and in the sarcin/ricin loop [30] and all of their spectroscopic 'trademarks' are reproduced here, including the upfield shift of the anomeric proton resonance belonging to the G of the GCs immediately adjacent to the AGs. Remarkably, only 11 of the 42 nucleotides of AD3 make base–base interactions that cannot be identified qualitatively by inspection of its spectra: the six bases in the middle of loop E, U74–G76 and G100–G102, and the five bases in loop D, G86–C90.

Distance restraints

Quantitative interpretation of the NMR data was required to establish the conformations of the less obvious parts of AD3. For the purposes of structure calculation, NOEs were classified as strong, medium, or weak based on their relative intensities in short mixing time NOESY spectra and given distance bounds of 2.4 ± 0.6 Å, 3.0 ± 1.0 Å, or 4.0 ± 1.0 Å, respectively.

A large number of distance restraints involving exchangeable protons were derived from the H_2O-NOESY spectra of AD3. Because exchange phenomena complicate the interpretation of such cross-peaks, wide bounds were assigned to protons correlated this way (2.5–5.0 Å). H_2O-NOESY spectra collected at short mixing times demonstrated that many of the imino-upfield NOEs observed in spectra obtained at longer mixing times result from magnetization transfers in which exocyclic amino protons act as intermediaries. Considerable effort was made to assign as many amino proton resonances as possible, and to distinguish direct imino-upfield NOEs from indirect imino-upfield NOEs, as well as to identify imino–amino NOEs. The chemical identities of amino proton resonances were determined using a ^{15}N-^{1}H HSQC optimized for amino proton couplings (data not shown). NOEs involving amino proton resonances could then be assigned specifically in a WATERGATE-NOESY experiment [31]. The correct base pairing in the center of loop E would not have emerged if these steps had not been taken.

A total of 254 internucleotide and 121 intranucleotide NOE-derived distance restraints were included in the calculations described below (Table 1). In addition, an unorthodox class of distance restraints was used, which we call 'unoes', to keep pairs of protons apart that do not give NOEs [32]. Unoes were included in structure calculations when structures that were otherwise acceptable were found to place pairs of protons so close together that NOEs would be expected that were not observed. Canonical hydrogen-bond distances with bounds of ± 0.1 Å were assigned to base pairs known to be Watson–Crick or GU wobble base pairs. We also included hydrogen-bond restraints for the type II, sheared A–G base pairs, A104–G72 and A78–G98, and the reversed Hoogsteen U–A pairs, U103–A73 and U77–A99.

Dihedral angle restraints

All the information available indicates that residues G68–C71, G105–C108, G79–G85 and C91–C97 in AD3 are located in double helices, and for this reason all of the torsion angles associated with these residues were restrained to be within $\pm 30°$ of ideal A-form values. As mentioned above, there were only two ^{31}P resonances

whose chemical shifts were outside of A-form range [24]: those belonging to G86-P-U87 and C88-P-U89. The α and ζ torsion angles of these two phosphates were left unrestrained. The α and ζ dihedral angles of all other residues were restrained to be ± 30° of A-form values.

Within loops D and E, the dihedral angle ε was restricted only from being *gauche*+, which is sterically forbidden [33,34]. The intraresidue H6-H5′ and H6-H5″ NOEs of U87 were both weak and of equal intensity, indicating that its γ-dihedral angle is in the A-form range. All other γ-dihedral angles in loops D and E were left unrestrained because their aromatic–H5′, H5″ NOEs could not be resolved. The β-dihedral angles of U87 and U89 were determined to be *trans* because there were no strong P-H5′ and H5″ couplings in the ^{1}H-^{31}P-COSY spectrum [27]. The other β-torsion angles in loop D and all β-torsion angles in loop E could not be assessed, and so were left unrestrained in the structure calculations. In addition, weak planarity restraints were included for bases that were found to form base pairs.

Table 1

Restraints for structure calculation.

Restraint	Number
Total NOEs	375
strong 2.4 ± 0.6 Å	83
medium 3.0 ± 1.0 Å	25
weak 4.0 ± 1.0 Å	266
internucleotide NOEs	254
intranucleotide NOEs*	121
Hydrogen bonds	40
Unoes	23
loop E	7
loop D	16
Dihedral angle restraints	406
backbone (α, β, γ, ε, ζ)	160
ribose pucker (ν_0–ν_4)	205
glycosidic (χ)	41
NOEs per residue	9.15
NOE and dihedral restraints per residue	19.05

*Only base to ribose NOEs are included.

Structure calculation

Structures were computed using a simulated annealing algorithm that permits only torsion angles to vary [32,35,36]. This procedure called torsion angle molecular dynamics (TAMD) starts with an extended nucleic acid chain with arbitrary dihedral angles that has the sequence of AD3. In each trial, it assigns rotational velocities to the variable torsion angles of the template at random, and then attempts to compute a structure consistent with the constraints supplied using molecular dynamics. About 20% of the trials yielded structures that had no NOE or dihedral angle violations greater than 0.5 Å or 5°, respectively.

Structures meeting this acceptance criterion were subsequently refined using a Cartesian molecular dynamics protocol in which electrostatics and attractive van der Waals potential energy terms were gradually turned on. An average structure was calculated from the nine refined structures that emerged, by averaging the Cartesian coordinates of each atom and then reminimizing its energy to ensure proper geometry. The minimized average structure is much closer to the initial average structure than any one of the nine structures in the set used to calculate it. The root mean square deviation (rmsd) between the initial averaged set of coordinates and the minimized average was 0.265 Å. Figure 5a is a stereo diagram of the minimized average structure of AD3 RNA, and Figure 5b shows the average structure in a CPK, space-filling representation.

The average pairwise, all-residue rmsd of each refined structure to the minimized average (residues 68–108) was 1.63 Å, but local features of the molecule superimposed much more closely. For loop E (residues 72–78, 98–104) and loop D (residues 86–90), the average pairwise rmsds were 0.90 Å and 0.55 Å, respectively (Figures 5c and 5d). Locally, the structure of AD3 is well defined by the NMR data; the relationship between its more distant parts is less well defined.

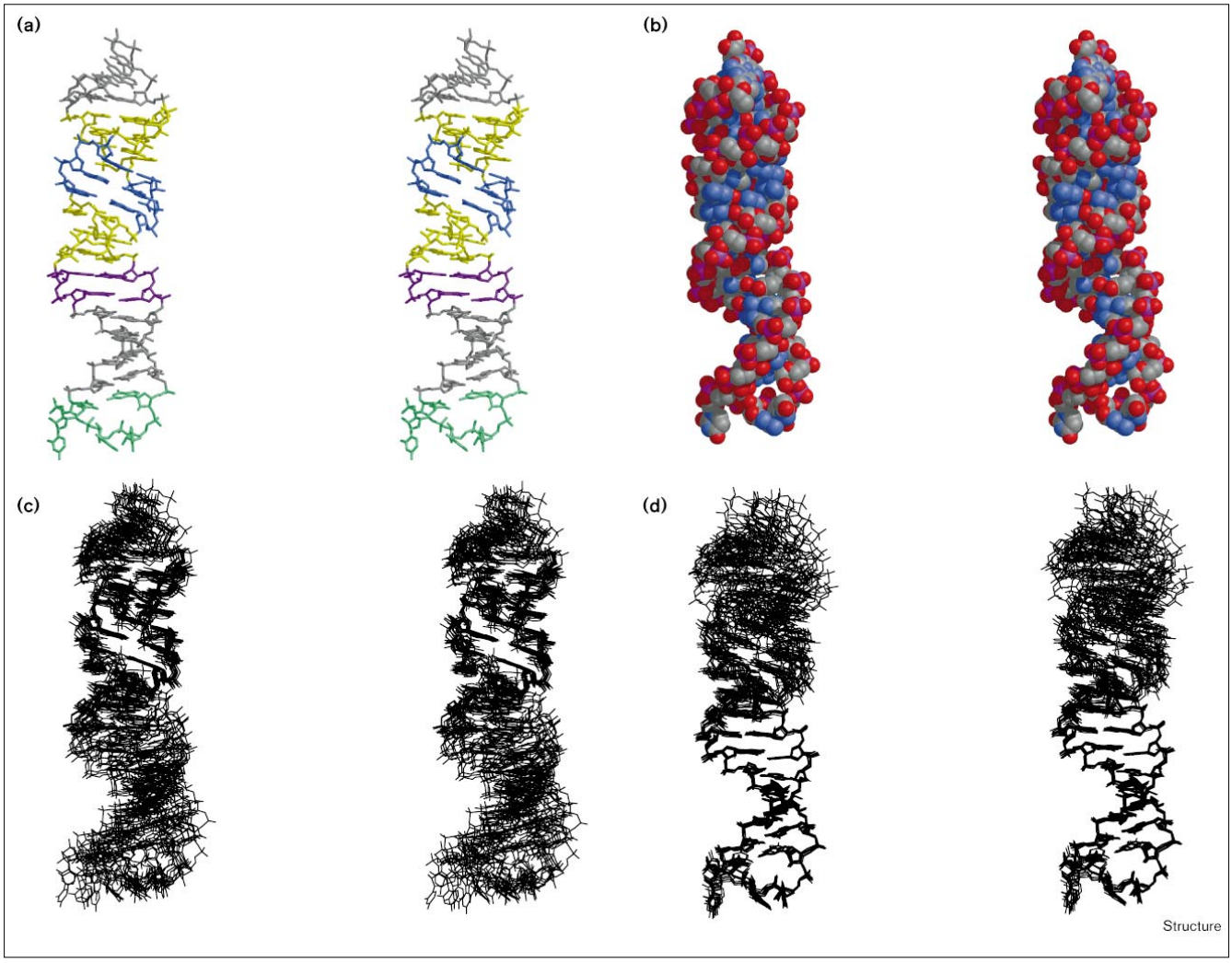
Validations of structure calculations

Because several of the restraints used in the structure calculations were inferred from the NMR spectra rather than obtained by direct measurement, test calculations were performed to assess their impact on the structure obtained. Four different restraints were tested: α and ζ dihedral angle restraints based on ^{31}P chemical shifts [24]; unoe restraints; the use of hydrogen-bond restraints in the palindromic ends of loop E; and the planarity restraints applied to the central base pairs in loop E.

A family of structures was computed for an 18-mer containing only loop E residues, in which α and ζ were unrestrained for all the noncanonically paired residues. In these structures, the α and ζ angles that emerged were the same as in the original calculation. The average structure of this test set superimposed on the original average structure with a 1.45 Å rmsd, indicating that no major changes in the organization of loop E had occurred. A similar calculation was performed on a 19-mer that contained only helix IV–loop D residues. The backbone torsion angles that emerged implied the same rotamers for each residue as in the original computations.

We also tested the convergence of structures calculated without the use of unoes. For loop D, 27 of the 31 zero-violation structures obtained had the same overall topology as

Figure 5

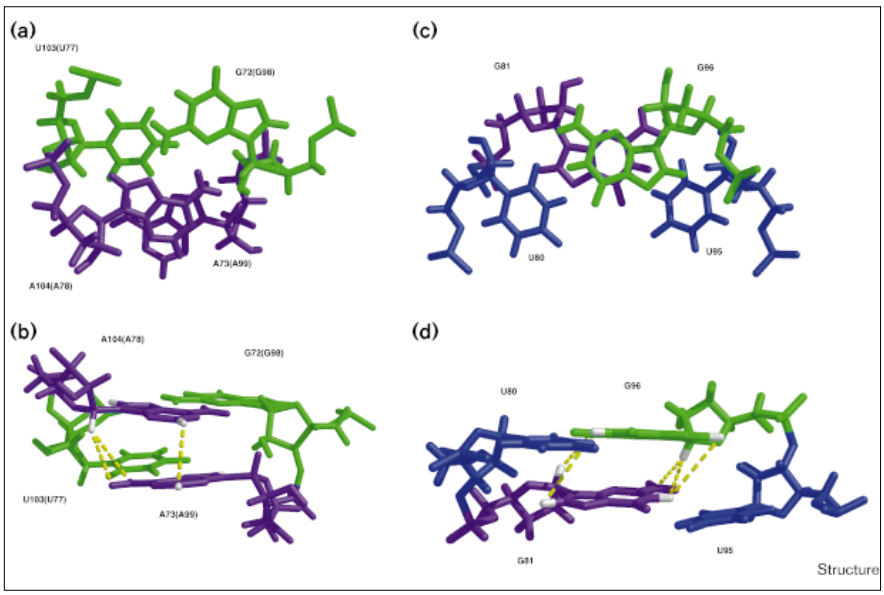


Stereo diagrams of the average structure of AD3 RNA. **(a)** The average structure of AD3 with regions of noncanonical structure elements shaded according to the color-coding in Figure 1c: helices IV and V are colored gray, the palindromic ends of loop E are yellow, the central base pairs of loop E are dark blue, the tandem GU wobbles purple, and loop D is colored green. **(b)** A heavy-atom only, CPK, space-filling representation of the average structure of AD3. Atoms are colored by type (nitrogen in blue, oxygen in red, phosphorus in magenta and carbon in gray). Superposition of **(c)** the loop E residues (72–78, 98–104) and **(d)** loop D–helix IV residues (79–97) from the nine structures generated by TAMD on the average structure. The average pairwise root mean square deviation between each structure and the average for loop E residues and loop D–helix IV residues is 0.90 Å and 0.55 Å, respectively.

determined when the unoes were included in the restraint set. The remaining four zero-violation structures had a higher energy than the majority class, and all of them place G86 in a position that implies the existence of NOEs that are not observed.

When planarity restraints were omitted for the central pairs of loop E (residues 74–76, 100–102), the structures that emerged exhibited the same base-pairing interactions as when the restraints were included, but the base pairs in some of the structures had a slightly buckled appearance. Overall, the average structure of this test set superimposed on the original average structure with a 1.23 Å rmsd. When the hydrogen-bond distances were omitted for the sheared AGs and the reversed Hoogsteen UAs in loop E, the hydrogen-bonding interactions implied by the structures obtained were identical to those specified originally. While the spread of structures in this test set was a bit higher than when the hydrogen bond restraints were included in the computation (1.21 Å average pairwise rmsd to the calculated average versus 0.90 Å for the original set), the computed average for the set without hydrogen-bond restraints superimposed on the original average with a 0.61 Å rmsd.

Figure 6



Cross-strand stacking motifs in AD3. View down the helix axis **(a)** and side view **(b)** of the palindromic motif(s) in loop E. The base pairing and stacking interactions are shown for the residues in the palindromic motifs on either side of the loop E motifs (residues 77–78, 98–99 and 72–73, 103–104). The adenine residues (colored purple) A104 and A73 (A78 and A99), form a cross-stranded stacking interaction. Dashed yellow lines between protons highlight NOEs between the stacked adenine residues. View down the helix axis **(c)** and side view **(d)** of the cross-strand stacked residues G96 (green) and G81 (purple) in the tandem GU wobbles at the start of helix IV. The six-membered rings of G81 and G96 are stacked on top of one another. The GU base pairs are shown with several NOEs that are crucial for defining this interaction, illustrated by dashed yellow lines. Uridines are shown in blue.

Conformation of loop E

Loop E has a partially palindromic sequence, which suggests that its structure might have approximate twofold symmetry (Figure 1c), which is indeed the case. The patterns of NOEs generated by the symmetry-related parts of the sequence are similar, as are the chemical shifts of the corresponding nuclei. This is evident in Figure 4b, in which the $^{13}C/^{1}H$ cross-peaks for most of the twofold related H1s in loop E are highlighted in color. Thus the spectroscopic data gave one reason to believe that the (GC)(AG)(UA) sequences at the two ends of loop E would be similar and, in the average structure, the palindromic ends of the loop superimpose on each other with a 0.37 Å rmsd.

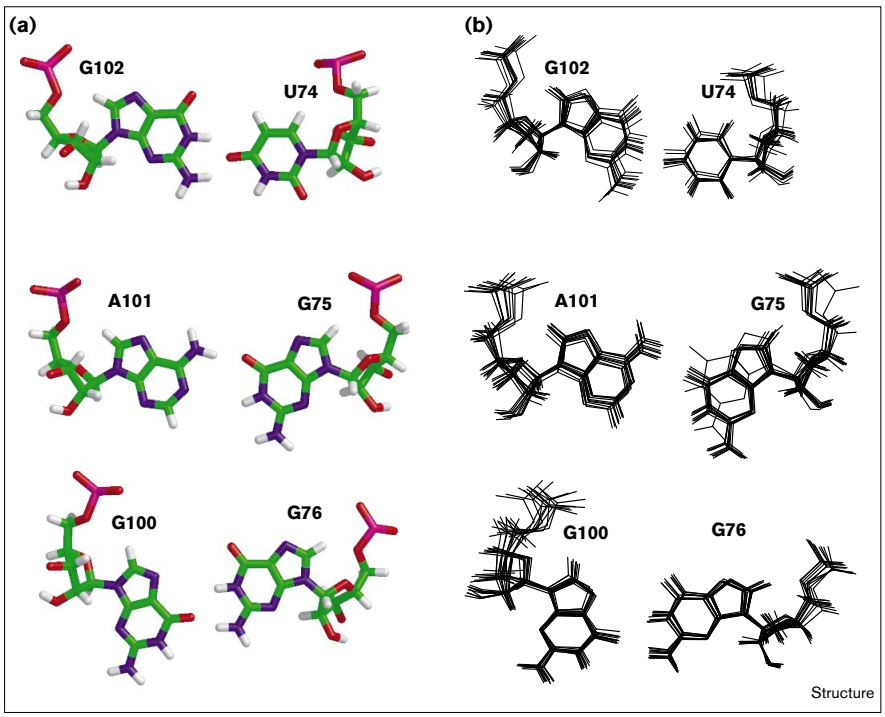
A Watson–Crick G–C base pair closes the terminal palindromic motifs at both ends of loop E. This is followed by a type II (*anti–anti*) sheared A–G base pair. The chemical shifts of the imino protons of G98 and G72 are 10.5 and 10.6 ppm, respectively, which is consistent with pairings in which imino protons are not involved in hydrogen bonds. The following base pair on either side of the loop is a reversed Hoogsteen U–A base pair. In this case, the chemical shifts of the imino protons of U103 and U77 are consistent with direct base–base hydrogen bonding, and strong NOEs from U77 H3 (U103 H3) and A99 H8 (A73 H8) identify these pairs. In the average structure A78 H62 (A104 H62) is within hydrogen-bonding distance (2.0 Å) of G98 O2′ (G72 O2′) which provides additional stabilization for the structure.

In these palindromic motifs, the adenine bases in the AG base pairs, A104 and A78, stack on the adenine bases of the reversed Hoogsteen pairs adjacent to them, A73 and A99 (Figure 6a). This cross-strand stacking is supported by several NOEs including a strong NOE between the H2 protons of A73 and A104 and also of A99 and A78 (Figure 6b). There is a kink in the backbone at residues A73 and A99. The β and γ angles are in the *gauche*$^{+}$ and *trans* conformer, respectively, instead of their A-form rotamers (*trans* and *gauche*$^{+}$, respectively). This kink produces a narrowing of the major groove at the center of loop E.

Each of the residues at the center of loop E (74–76, 100–102) forms a noncanonical pair with the residue directly opposite it in the sequence. Figure 7a shows the pairings in the average structure, and Figure 7b shows the superpositions of these pairs from each of the final structures onto the average pair.

We now have much more information about the G102–U74 pair than we did in 1995 [10]. It has long been clear that this pair is not a wobble GU; the most distinctive spectroscopic feature of a GU wobble pair is a very strong NOE between its G and U imino protons. In the spectra of AD3, only the weakest of cross-peaks is observed between the G and U imino protons of this pair, and it is only visible at 5°C. Furthermore, the chemical shift of the imino of U74 (10.7 ppm) is about 1 ppm upfield of where UH3 protons resonate in wobble GUs, and an NOE is observed between G102 H1 and U74 H5, which is not seen in wobble GUs. The pairing that emerged from our calculation implies the existence of a bifurcated hydrogen bond, between the carbonyl oxygen of U74 O4 and G102 H1 and G102 H21. The average distance between U74 O4 and G102 H1 in the set of structures calculated was 1.46 Å and the standard deviation is

Figure 7



Base pairings in the center of loop E. **(a)** The pairings for G102–U74 (top), A101–G75 (middle) and G100–G76 (bottom) that resulted from the average structure calculation. **(b)** Superpositions of individual base pairs in loop E. The average pairwise root mean square deviations for the family of calculated structures are 0.30 Å, 0.43 Å and 0.41 Å for G102–U74 (top), A101–G75 (middle) and G100–G76 (bottom), respectively.

0.07 Å. For U74 O4 and G102 H1, the distance was 2.20 ± 0.26 Å. This is the base pair that has changed the most since the description of the initial model. In the 1995 model, the pairing proposed for U74–G102 depended on a hydrogen bond between U74 O4 and G102 H22, which is inconsistent with chemical reactivity data [37]. The current pairing, like that of the 1995 model, explains why loop E is destabilized when cytosine replaces U74 [38], but it is also consistent with the chemical reactivity data obtained by the Ehresmann group [37]. The average pairwise rmsd of the superposition of the G102–U74 pairs in the final set of structures with that in the average structure is 0.30 ± 0.18 Å.

The G75–A101 pairing is the one that is the least constrained by our data. Nevertheless, in all accepted structures it forms an opened up, Watson–Crick-like A–G pair that is stabilized by only a single hydrogen bond: the distance between G75 O6 and A101 H61 is 1.95 ± 0.11 Å. The imino proton of G75 is oriented toward the N1 of A101, but the distance is too long for a hydrogen bond (3.79 ± 0.11 Å), which may explain why we do not observe a G75 H1 resonance. In our preliminary model the pairing proposed for G75–A101 included this hydrogen bond, but that proposal depended on an imino proton resonance assignment we have since had to change. The average pairwise rmsd of the superposition of the G75–A101 pairs in the final set of structures with that in the average structure is 0.43 ± 0.17 Å.

In our preliminary model for loop E, G76 was assumed to make a reversed Hoogsteen-like pair with G100 but, as we pointed out at the time, this did not fully explain the data [10]. If the two bases were paired this way, an NOE should be observed between G76 H1 and G100 H8, and it is not. In addition, G76 H1 should resonate well downfield, which it also does not. The pairing that emerged from these computations explains both discrepancies (Figure 7; bottom). The G76–G100 pair is stabilized by a bifurcated hydrogen bond between G100 O6 and both G76 H1 (2.20 ± 0.05 Å) and G76 H21 (1.85 ± 0.07 Å). The upfield chemical shift of G76 H1 (10.2 ppm) is explained by its failure to hydrogen bond to any of its partner's ring nitrogens, and the distance between it and the H8 of G100 is quite large (5.0 Å). The average pairwise rmsd for the superposition of the G76–G100 pairs in the set of final structures with that in the average structure is 0.41 ± 0.05 Å.

Table 2 shows the dihedral angles of all loop E residues and their standard deviations. It should be noted that in every case where more than one rotamer was populated within the family of structures, the average structure has the rotamer that is populated in the majority of the structures that were averaged.

Cross-strand stacking in successive GU wobble base pairs

Directly adjacent to the closing G79–C97 base pair of loop E towards helix IV, there are two successive GU base pairs

Table 2

Average values of dihedral angles in the nine calculated structures of AD3 ± standard deviations.

Residue	c	α	β	γ	δ	ε	ζ
Loop E							
G72	197 ± 3	292 ± 1	**157 ± 1**	**66 ± 6**	83.0 ± 2	**206 ± 22**	265 ± 2
A73	168 ± 5	270 ± 22	**67 ± 19 (7)** **135 ± 0 (1)** **312 ± 0 (1)**	**171 ± 27**	93 ± 3	**235 ± 5**	317 ± 7
U74	193 ± 4	288 ± 6	**188 ± 8**	**41 ± 6**	80 ± 1	**221 ± 7**	293 ± 5
G75	187 ± 4	289 ± 9	**152 ± 8**	**73 ± 10**	81 ± 2	**232 ± 5**	297 ± 4
G76	173 ± 5	266 ± 2	**140 ± 11**	**101 ± 12**	82 ± 3	**210 ± 7**	295 ± 4
U77	196 ± 2	281 ± 11	**193 ± 6**	**53 ± 12**	85 ± 2	**225 ± 12**	289 ± 6
A78	177 ± 1	271 ± 7	**146 ± 10**	**93 ± 5**	80 ± 1	**219 ± 4**	308 ± 5
G98	205 ± 3	292 ± 4	**172 ± 3**	**47 ± 3**	83 ± 2	**192 ± 25**	267 ± 4
A99	176 ± 11	281± 27	**60 ± 3 (6)** **199 ± 1 (3)**	**169 ± 5 (6)** **53 ± 6 (3)**	90 ± 10	**257 ± 16**	303 ± 11
G100	176 ± 8	278 ± 16	**128 ± 30**	**99 ± 34**	80 ± 3	**228 ± 5**	301 ± 4
A101	194 ± 3	283 ± 7	**169 ± 13**	**58 ± 13**	84 ± 3	**216 ± 4**	293 ± 4
G102	178 ± 2	273 ± 11	**151 ± 7**	**89 ± 14**	80 ± 2	**211 ± 4**	295 ± 7
U103	196 ± 3	293 ± 9	**187 ± 10**	**45 ± 6**	82 ± 2	**218 ± 9**	288 ± 4
A104	187 ± 2	292 ± 10	**154 ± 6**	**73 ± 8**	79 ± 2	**226 ± 4**	297 ± 4
Loop D							
U87	248 ± 2	**170 ± 5**	155 ± 0	70 ± 3	87 ± 0	**201 ± 2**	**293 ± 1**
C88	203 ± 2	265 ± 2	**209 ± 2**	**17 ± 1**	146 ± 2	**278 ± 3**	**182 ± 8**
U89	230 ± 8	**88 ± 3 (3)** **176 ± 2 (5)** **335 ± 0 (1)**	162 ± 25	**40 ± 1 (5)** **157 ± 48 (4)**	82 ± 2	**177 ± 6**	**82 ± 3**

Multiple values are shown for a given dihedral angle where different rotamers were populated within the family of accepted structures. In these cases, the number in parentheses reflects the number of structures within the set that adapted that particular conformation. The values of the dihedral angles that are in boldface are the angles that were left unrestrained in the structure calculation.

in the orientation (UG)(GU). Spectroscopically, these base pairs exhibit the standard wobble GU trademarks discussed above. An interesting feature of these two GU wobbles, which was not anticipated, is that the guanine bases that are on opposite strands are stacked in a cross-stranded fashion (Figure 6c). Specifically, the six-membered rings of G81 and G96 stack directly on top of one another. NOEs that define this interaction involve NOEs between G96 H1 and G81 H8 as well as between G81 H1 and G96 H8 (Figure 6d). The average pairwise rmsd of these residues with the calculated average is 0.24 ± 0.07 Å.

The conformation of loop D

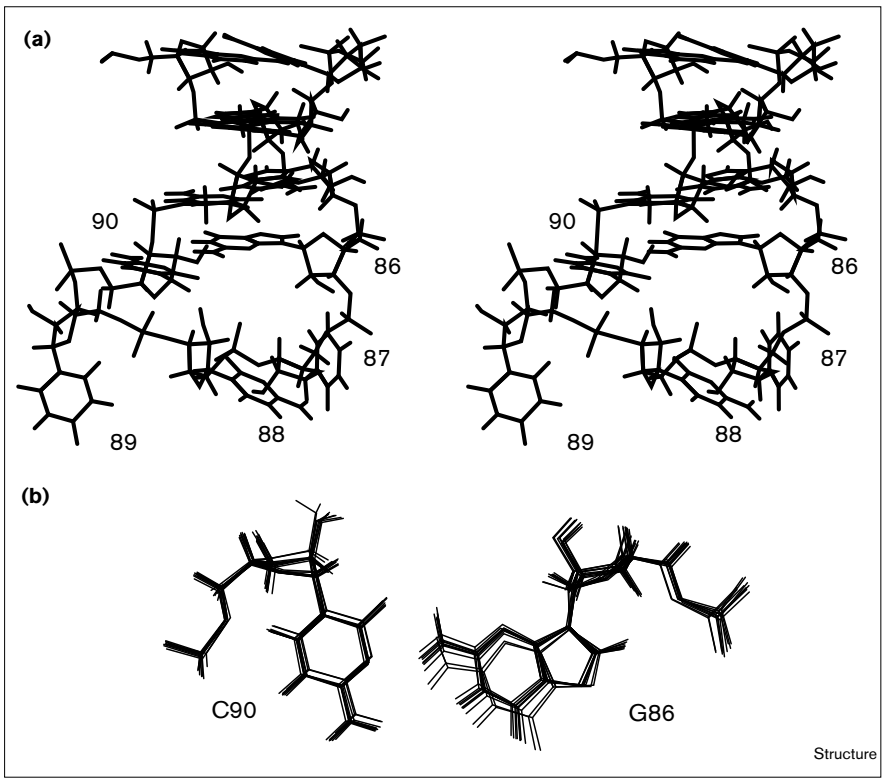
Loop D is an all-pyrimidine hairpin loop. The structure of this loop must be more mobile relative to the molecule as a whole because the ^{1}H and ^{13}C nuclei from residues U87, C88 and U89 have the longest transverse relaxation times of all residues in the molecule (AD and PBM, unpublished results; [10]). Nevertheless, a single average structure emerged for loop D (Figure 8a).

No imino proton resonances are observed in loop D or in its closing base pair, G86–C90. For this reason, the closing G–C base pair was not restrained to be a Watson–Crick pair during structure calculations. The G–C pairing that emerged has the H21 amino proton of G86 within hydrogen-bonding distance (2.16 ± 0.16 Å) of C90 O2 (Figure 8b), but the orientations of the hydrogen-bond donor and acceptor are unusual enough to give one pause. Starting at the 5′-side of the loop, residues U87 and C88 are stacked and are oriented with their bases pointing towards the minor groove, and there are many base-to-base NOEs that support this stacking interaction (data not shown). U89 is then flipped out into solution. The phosphates associated with residues U87 and U89 have ^{31}P chemical shifts that are not in A-form range. Table 2 shows that the dihedral angles α and/or ζ in all of the structures computed are, indeed, out of A-form range for residues U87 and U89.

Discussion

Structure calculations were done on AD3 using both distance geometry and TAMD methods. In our hands, only the TAMD protocol produced acceptable structures for the whole molecule. Distance geometry/simulated annealing (DGSA) computations run with the same set of restraints produced no zero-violation structures. However, DGSA calculations done on an 18-mer that included only loop E residues and its closing G–C base pairs, and on a 19-mer that included all of helix IV and loop D (79–97) did produce families of structures that met the same acceptance criteria as those generated using TAMD methods, but there was a big difference in yield. Of the

Figure 8



The solution structure of loop D. **(a)** Stereo drawing of the average NMR structure of loop D (residues 83–93) is shown. **(b)** Superposition of the closing base pairs (G86–C90) from the family of calculated structures on the average structure. The average pairwise root mean square deviation for the G86–C90 pair in the family of calculated structures is 0.30 Å.

TAMD trials run on the loop E region alone (residues 71–79, 97–105) or the helix IV–loop D region alone (residues 79–97), 80–95% were successful. Only 15–20% of the DGSA trials run on the same sequences with the same restraint sets succeeded. The rmsds of the families of structures generated by the two methods indicate that the spread of computed structures did not differ significantly with either of the two computation protocols (0.97 Å for DGSA, 0.99 Å for TAMD).

The structure we have obtained for the loop E–loop D arm of *E. coli* 5S rRNA includes some familiar RNA motifs as well as several novel features. Loop E is closed on either side by a motif (AG)(UA) that has been found in several other structures: the sarcin–ricin loop [30] and eukaryotic 5S rRNA loop E [29]. Despite additional sequence similarities, the conformational similarities between the *E. coli* loop E and these other structures are limited to these four bases. In both the sarcin–ricin and eukaryotic loop E, the cross-strand stacking motif sits directly above a bulged guanine. The structure of prokaryotic loop E does not include any bulged residues even though there are guanines on either side of these motifs in the appropriate position. Instead, the central six residues of loop E in *E. coli* form three irregular base pairs, each with the residue directly across from it in sequence. The backbone of loop E is distorted from standard A-form geometry in the palindromic motifs at both ends of the loop, which may explain why the base pairs at the center of loop E have such unconventional geometries. Because of the distortion in the backbone at the ends of loop E, the cross-strand (e.g., C1′–C1′) distances in the central three base pairs are 2–4 Å longer than the corresponding distances in Watson–Crick double helix regions. It is possible that the more 'open' base pairs found at the center of loop E (especially G75–A101) are formed in response to the unusually large cross-strand distances that must be spanned because of the unusually large backbone separations.

The structure obtained for the palindromic motifs in loop E is consistent with chemical probing experiments that have been carried out on *E. coli* 5S rRNA [37]. The N1 faces of all the adenine residues in these motifs are reactive to dimethyl sulfate (DMS), and in our structure they are exposed to solvent. In addition, the N1 faces of G72 and G98 in the A–G pairs do not react as is the case in sheared A–G base pairs in other molecules. The chemical protection patterns for the central pairs in loop E can also be rationalized. The N1 faces of G75, G76 and G102 are all unreactive and, as Figure 7a shows, in the case of G76 and G102 the H1 protons are involved in bifurcated hydrogen bonds. The imino proton of G75 is not directly hydrogen bonded to A101, but it does point towards A101, which could explain its resistance to chemical modification.

The conformation of loop D is also consistent with the results of biochemical experiments. The chemical probing data of Brunel, *et al.* [37], indicate that all of the bases in loop D are reactive to a variety of chemicals including cyclohexyl-N′-[2-(N-methylmorpholino) ethyl] carbodiimide *p*-toluene sulfate (CMCT), DMS, Pb^{2+}, and in our structure they are exposed. It is interesting that the base most exposed to solution, U89, is the one residue that cross-links to sites on 23S rRNA near the peptidyl transferase ring [5,6].

It is important to stress that the spectroscopic data do not provide direct evidence for the nature of the pairing of G86 and C90. There are no NOEs that eliminate the possibility of Watson–Crick pairing, or that specifically corroborate the pairings observed in the structures we calculated. As this base pair closes a loop, it might be more exposed to solvent than normal. This alone could account for our failure to observe a G86 H1 imino proton resonance, even if it were a Watson–Crick pair. In fact, when Watson–Crick hydrogen-bonding distances were imposed on this base pair in a test calculation, the zero-violation low energy structures that emerged had the same topology as the loop shown here. U89 was still extruded into solution and the bases of U87 and C88 were still stacked, and pointed towards the minor groove.

The crystal structure of a dodecamer that includes the loop E region of *E. coli* 5S rRNA has recently been determined at 1.5 Å resolution, as has the crystal structure of fragment 1 at 3 Å resolution [39]. These crystal structures and the NMR structure presented above agree. Specifically, the base pairings in loop E are the same in both structures, and the heavy atoms in loop E superimpose with a rmsd of 1.3 Å. The loop E–loop D arm of *E. coli* 5S rRNA is the first case we are aware of where a crystal structure and a solution structure of an RNA have become available for closely related sequences. These structures are being compared in detail at this time, and the results of that analysis will be published elsewhere.

Biological implications

5S rRNA is a component of the large (50S) subunit of the prokaryotic ribosome. The solution structure of the 5S rRNA domain that contains loops E and D has been determined using both homonuclear and heteronuclear multidimensional NMR spectroscopy. It has long been known that the binding site for the ribosomal protein L25 on 5S rRNA includes loop E and that L25 binding stabilizes the structure of this loop [7,40,41]. The solution structure reveals that while the loop E–loop D region is double helical over its entire length, the geometry of that helix is quite abnormal around loop E and the adjoining pair of GU base pairs. The major groove of loop E is narrow, while the minor groove is broad and almost flat. Furthermore, the array of hydrogen-bond donors and acceptors presented on the surface of the minor groove is very different from those that would be present if loop E were made up of Watson–Crick base pairs. In the region of the two GU wobble pairs, on the other hand, the major groove is much wider than in regular A-form helix, so wide in fact that an α helix could enter it. It is likely that the sequence-specific interaction of L25 with this part of 5S rRNA depends on these unusual structural features.

Except for the reversal of one of its terminal GCs, the sequence of loop E in *E. coli* 5S rRNA is the same as the consensus sequence for eubacterial 5S rRNAs (RR Gutell, personal communication). Thus, it is likely that the loop E regions of other prokaryotic 5S rRNAs will resemble the one shown here. The solution structure of eukaryotic 5S rRNA loop E has also been determined and is quite different [29]. The sequence of loop E is not palindromic in eukaryotes, and the structure has no twofold character. The structure of eukaryotic loop E, however, is not entirely unrelated to that of the prokaryotic loop E: eukaryotic loop Es 'begin' with a single copy of the (AG)(AU) cross-strand purine stack motif that is capped with a Watson–Crick GC. Why prokaryotic loop Es contain two such motifs and eukaryotic loop Es contain one is unclear but, at least in part, this structural difference may explain why eukaryotic 5S rRNAs do not substitute for eubacterial 5S rRNAs in large ribosomal subunit reconstitutions [42].

The (AG)(AU) motif that appears twice in AD3 has been found in other RNAs [29,30]. There is much similarity between the geometry of these motifs and that of the (GU)(UG) juxtaposition found in the 5S rRNA helix IV. In both types of motif, purines in opposite strands stack on top of each other rather than stacking on their neighbors in the same strand. Thus both are examples of cross-strand purine stacks, which are likely to be very common in RNA. A recently published crystal structure of a synthetic RNA containing the same (GU)(UG) sequence flanked on both sides by Watson–Crick helix has the same conformation as ours [43].

One of the most striking features of AD3s structure are the three base pairs in the middle of loop E, which are stabilized either by single base–base hydrogen bonds or by single, bifurcated base–base hydrogen bonds. An important lesson this structure teaches is that the universe of base pairings possible in RNAs is significantly larger than the 28 two hydrogen bond pairings that have long been considered standard [44].

Materials and methods

NMR sample preparation

AD3 RNA was synthesized by *in vitro* transcription using T7 RNA polymerase and the run-off plasmid template pAD3 as described in [10], with the exception that the plasmid was purified from cultures using a QIAGEN Plasmid GIGA kit [45–47]. Unlabeled NTPs used in

the transcriptions were purchased from Pharmacia. Uniformly-labeled $^{13}C/^{15}N$-NMPs were prepared from *E. coli MRE600* grown on minimal media containing ^{13}C-labeled glucose and ^{15}N-labeled NH_4Cl and enzymatically phosphorylated to produce triphosphates according to the protocols described in [48,49]. AD3 RNA was purified from transcription reactions as described in [10] and was then dialyzed extensively against NMR buffer (100 mM KCl, 5 mM cacodylate, 4 mM $CaCl_2$, and 0.1 mM EDTA, pH 7.2 or pH 5.5). The RNA was concentrated to its final volume of 180–200 µl in Centricon-3 concentrators (Amicon). ^2H_2O was added to samples dissolved in H_2O for a spectrometer lock, and dioxane was added to all samples as an internal chemical shift standard. For experiments carried out in ^2H_2O, the NMR sample was lyophilized and resuspended in 99.5% ^2H_2O three times. The sample was then lyophilized one more time and resuspended in 99.96% ^2H_2O. Shigemi NMR microtubes were used to contain the prepared samples. The final concentrations of RNA ranged from 1.0–3.0 mM.

NMR spectroscopy

Proton and heteronuclear NMR spectra were collected either on a Varian Unity v500 spectrometer or a Varian Unity+ v600 spectrometer, both equipped with triple resonance probes with z axis gradients, or on a GE Ω500 spectrometer with a triple resonance gradient probe equipped with triple-axis gradients. ^{31}P spectra were collected on a GE Ω500 spectrometer with a triple resonance probe with ^1H, ^{13}C and ^{31}P channels and z axis gradients. ^1H and ^{13}C chemical shifts were referenced to dioxane at 3.741 ppm and 69.2 ppm, respectively. The carrier frequency was set on H_2O for ^1H experiments. The ^{13}C carriers were set to 77.5 or 123 ppm depending on whether the experiment was optimized for ribose or aromatic couplings. The ^{15}N carrier was set to 153 ppm or 85 ppm for imino- or amino-optimized experiments, respectively.

Spectra in H_2O were collected using either gradient-enhanced, jump-return, spin-echo [50] or WATERGATE water suppression [31]. NOESY spectra in ^2H_2O were collected with a pre-delay of 6.0 s and mixing times of 35, 50, 75, 100 and 150 ms. All other 2D and 3D experiments that involved ^{13}C or ^{31}P were implemented in a manner similar to that described in Stallings and Moore [32].

Restraints for the structure calculation

Distances between protons were estimated using the relative intensities of cross-peaks in NOESY spectra. Cross-peaks that were conspicuous in a 35 ms NOESY were classified as being strong. The protons correlated by such NOEs were assigned a distance of 2.4 ± 0.6 Å. NOEs that appeared weakly in the 35 ms NOESY spectrum were classified as medium and their protons were restrained to be 3.0 ± 1.0 Å apart. Cross-peaks that appeared only in 100 ms or 150 ms NOESYs were considered weak, and their protons were assigned distances of 4.0 ± 1.0 Å. Hydrogen bond distance restraints were imposed on residues whose base pairing could be inferred from the NMR data. This includes all canonical base pairs in helix IV and helix V including the GU wobble pairs, as well as the sheared AG and reversed Hoogsteen base pairs at the ends of loop E. Weak planarity restraints were used (50 kcal mol^{-1} $Å^{-2}$) for residues that were found to be base paired. In addition, distance restraints that we call unoes, that forced two protons that did not have an NOE between them to be more than 5 Å apart (note that a lower bound distance of 4.5 Å is permitted within the acceptance criteria) were included when structures were repeatedly calculated that predicted NOEs that were not observed. The dihedral restraints used are described in the text, and can be inferred from Table 2.

Structure calculation

Structures were calculated using the TAMD algorithm described in [32]. The parameter and topology sets used were a modified version of those published by the Berman group [51]. They include a restraint to ensure that the chirality of the non-bonded phosphate oxygens is maintained throughout the structure calculation (JP Rife, unpublished results). The protocol used was identical to that described in Stallings and Moore [32]. Structures that met the acceptance criteria were then minimized with a Cartesian space refinement protocol that included attractive van der Waals energy terms as well as electrostatics.

Graphics

Molecular graphics images were produced using the MidasPlus program from the Computer Graphics Laboratory, UCSF (supported by NIH RR-01081).

Accession numbers

The coordinates of the nine lowest energy structures and the computed average are being deposited into the Protein Data Bank at Brookhaven.

Supplementary material

Supplementary material available with the internet version of this paper contains a table listing all of the assigned proton chemical shifts for AD3.

Acknowledgements

We thank Axel Brünger for access to the TAMD protocol used in the structure calculation prior to its becoming publicly available and SC Stallings for technical assistance in the implementation of TAMD. This work was supported by grants from the NIH (AI-09167 prior to 12/95, and GM-41651 after 12/95).

References

1. Moore, P.B. (1995). Structure and function of 5S RNA. In *Ribosomal RNA: Structure, Evolution, Processing and Function in Protein Synthesis.* (Zimmermann, R.A. & Dahlberg, A.E., eds), pp. 199–236, CRC Press, Boca Raton, FL.
2. Shatsky, I.N., Evstafieva, A.G., Bystrova, T.F., Bogdanov, A.A. & Vasiliev, V.D. (1980). Topography of RNA in the ribosome: location of the 3′ end of 5S RNA on the central protuberance of the 50S subunit. *FEBS Lett.* **121**, 97–100.
3. Selivanova, O.M., Gongadze, G.M., Gudkov, A.T. & Vasiliev, V.D. (1986). Structure of protein-deficient 50S ribosomal subunits: particles without 5S RNA–protein complex retain the L7/L12 stalk and associate with 30S subunits. *FEBS Lett.* **197**, 79–82.
4. Fox, G.E. & Woese, C.R. (1975). 5S RNA secondary structure. *Nature* **256**, 505–507.
5. Dontsova, O., *et al*., & Brimacombe, R. (1994). Stem-loop IV of 5S rRNA lies close to the peptidyltransferase center. *Proc. Natl. Acad. Sci. USA* **91**, 4125–4129.
6. Dokudovskaya, S., Dontsova, O., Shpanchenko, O., Bogdanov, A. & Brimacombe, R. (1996). Loop IV of 5S ribosomal RNA has contacts both to domain II and to domain V of the 23S RNA. *RNA* **2**, 146–152.
7. Douthwaite, S., Garrett, R.A., Wagner, R. & Feunteun, J. (1979). A ribonuclease-resistant region of 5S rRNA and its relation to the RNA binding sites of proteins L18 and L25. *Nucleic Acids Res.* **6**, 2453–2470.
8. Moore, P.B., Abo, S., Freeborn, B., Gewirth, D.T., Leontis, N.B. & Sun, G. (1988). Preparation of 5S RNA-related materials for nuclear magnetic resonance and crystallographic studies. *Methods Enzymol.* **164**, 158–174.
9. White, S.A., Nilges, M., Huang, A., Brünger, A.T. & Moore, P.B. (1992). NMR analysis of helix I from the 5S RNA of *Escherichia coli*. *Biochemistry* **31**, 1610–1620.
10. Dallas, A., Rycyna, R. & Moore, P. (1995). A proposal for the conformation of loop E in *Escherichia coli* 5S rRNA. *Biochem. Cell Biol.* **73**, 887–897.
11. Kime, M.J. & Moore, P.B. (1983). Physical evidence for a domain structure in *Escherichia coli* 5S RNA. *FEBS Lett.* **153**, 199–203.
12. Zhang, P. & Moore, P.B. (1989). An NMR study of the helix V-loop E region of the 5S RNA from *Escherichia coli*. *Biochemistry* **28**, 4607–4615.
13. Leontis, N.B., Ghosh, P. & Moore, P.B. (1986). Effect of magnesium ion on the structure of the 5S RNA from *Escherichia coli*. An imino proton magnetic resonance study of the helix I, IV, and V regions of the molecule. *Biochemistry* **25**, 7386–7392.
14. Szewczak, A.A., Kellogg, G.W. & Moore, P.B. (1993). Assignment of NH resonances in nucleic acids using natural abundance ^{15}N-^1H correlation spectroscopy with spin-echo gradient pulses. *FEBS Lett.* **327**, 261–264.

15. Gewirth, D.T. & Moore, P.B. (1987). Effects of mutation on the downfield proton nuclear magnetic resonance spectrum of the 5S RNA of *Escherichia coli*. *Biochemistry* **26**, 5657–5665.
16. Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*. John Wiley & Sons, New York, NY.
17. Vuister, G.W. & Bax, A. (1992). Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. *J. Magn. Reson.* **98**, 428–435.
18. Bax, A. & Sumners, M.F. (1986). ^{1}H and ^{13}C assignments from sensitivity-enhanced detection of heteronuclear multiple-bond connectivity by 2D multiple quantum NMR. *J. Am. Chem. Soc.* **108**, 2093–2094.
19. Marino, J.P., Prestegard, J.H. & Crothers, D.M. (1994). Correlation of adenine H2-H8 resonances in uniformly ^{13}C labeled RNAs by 2D HCCH-TOCSY: a new tool for ^{1}H assignment. *J. Am. Chem. Soc.* **116**, 2205–2206.
20. Nikonowicz, E.P. & Pardi, A. (1992). Simple procedure for resonance assignment of the sugar protons in ^{13}C-labeled RNAs. *J. Am. Chem. Soc.* **114**, 9202–9203.
21. Ikura, M. & Kay, L.E. (1991). Improved three-dimensional ^{1}H-^{13}C-^{1}H correlation spectroscopy of a ^{13}C-labeled protein using constant-time evolution. *J. Biomol. NMR* **1**, 299–304.
22. Nikonowicz, E.P. & Pardi, A. (1993). An efficient procedure for assignment of the proton, carbon, and nitrogen resonances in ^{13}C/^{15}N-labeled nucleic acids. *J. Mol. Biol.* **232**, 1141–1156.
23. Zhang, P., Rycyna, R. & Moore, P.B. (1989). A study of the conformation of 5S by ^{31}P NMR. *Nucleic Acids Res.* **17**, 7295–7302.
24. Gorenstein, D.G. (1981). Nucleotide conformational analysis by ^{31}P nuclear magnetic resonance spectroscopy. *Annu. Rev. Biophys. Bioeng.* **10**, 355–386.
25. Gorenstein, D.G. (1984). Phosphorus-31 chemical shifts: principles and empirical observations. In *Phosphorus-31 NMR: Principles and Applications* (Gorenstein, D.G., ed), pp. 7–36. Academic Press, Orlando, FL.
26. Sklenár, V., Miyashiro, H., Zon, G., Miles, H.T. & Bax, A. (1986). Assignment of the ^{31}P and ^{1}H resonances in oligonucleotides by two-dimensional NMR spectroscopy. *FEBS Lett.* **208**, 94–98.
27. Varani, G. & Tinoco, I., Jr. (1991). RNA structure and NMR spectroscopy. *Quart. Rev. Biophys.* **24**, 479–532.
28. Altona, C. (1982). Conformational analysis of nucleic acids. Determination of the backbone geometry of single-helical RNA and DNA in aqueous solution. *Recueil Review: Journal of the Royal Netherlands Chemical Society* **101**, 413–433.
29. Wimberly, B., Varani, G. & Tinoco, I., Jr. (1993). The conformation of loop E of eukaryotic 5S ribosomal RNA. *Biochemistry* **32**, 1078–1087.
30. Szewczak, A.A. & Moore, P.B. (1995). The sarcin/ricin loop, a modular RNA. *J. Mol. Biol.* **247**, 81–98.
31. Lippens, G., Dhalluin, C. & Wieruszeski, J.-M. (1995). Use of a water flip-back pulse in the homonuclear NOESY experiment. *J. Biomol. NMR* **5**, 327–331.
32. Stallings, S.C. & Moore, P.B. (1997). The structure of an essential splicing element: stem loop IIa from yeast U2 snRNA. *Structure* **5**, 1173–1185.
33. Blommers, M.J.J., Nanz, D. & Zerbe, O. (1994). Determination of the backbone torsion angle ε in nucleic acids. *J. Biomol. NMR* **4**, 595–601.
34. Lankhorst, P.P., Haasnoot, C.A.G., Erkelens, C. & Altona, C. (1984). Carbon-13 NMR in conformational analysis of nucleic acid fragments 2. A reparametrization of the Karplus equation for vicinal NMR coupling constants in CCOP and HCOP fragments. *J. Biomol. Struct. Dyn.* **1**, 1387–1405.
35. Rice, L.M. & Brünger, A.T. (1994). Torsion angle dynamics: reduce variable conformational sampling enhances crystallographic structure refinement. *Proteins* **19**, 277–290.
36. Stein, E.G., Rice, L.M. & Brünger, A.T. (1997). Torsion angle molecular dynamics: a new, efficient tool for NMR structure calculation. *J. Magn. Reson. B* **127**, 154–164.
37. Brunel, C., Romby, P., Westhof, E., Ehresmann, C. & Ehresmann, B. (1991). Three-dimensional model of *Escherichia coli* ribosomal 5S RNAs as deduced from structure probing in solution and computer modelling. *J. Mol. Biol.* **221**, 293–308.
38. Zhang, P. (1989). Studies of the helix V-loop E domain in the 5S RNA of *E. coli* by ^{1}H and ^{31}P NMR. PhD thesis, Yale University.
39. Correll, C.C., Freeborn, B., Moore, P.B. & Steitz, T.A. (1997). Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell*, in press.
40. Kime, M.J. & Moore, P.B. (1983). Nuclear Overhauser experiments at 500 MHz on the downfield proton spectra of 5S ribonucleic acid and its complex with ribosomal protein L25. *Biochemistry* **22**, 2622–2629.
41. Leontis, N.B. & Moore, P.B. (1986). Imino proton exchange in the 5S RNA of *Escherichia coli* and its complex with protein L25 at 490 MHz. *Biochemistry* **25**, 5736–5744.
42. Wrede, P. & Erdmann, V.A. (1973). Activities of *B. stearothermophilus* 50S ribosomes reconstituted with prokaryotic and eukaryotic 5S rRNA. *FEBS Lett.* **33**, 315–319.
43. Biswas, R., Wahl, M.C., Ban, C. & Sundaralingam, M. (1997). Crystal structure of an alternating octamer r(GUAUGUA)dC with adjacent GU wobble pairs. *J. Mol. Biol.* **267**, 1149–1156.
44. Saenger, W. (1984). *Principles of Nucleic Acid Structure*. Springer-Verlag, New York.
45. Szewczak, A.A., White, S.A., Gewirth, D.T. & Moore, P.B. (1990). On the use of T7 RNA polymerase transcripts for physical investigation. *Nucleic Acids Res.* **18**, 4139–4142.
46. Milligan, J.F., Groebe, D.R., Witheell, G.W. & Uhlenbeck, O.C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **15**, 8783–8798.
47. Wyatt, J.R., Chastain, M. & Puglisi, J.D. (1991). Synthesis and purification of large amounts of RNA oligonucleotides. *Biotechniques* **11**, 764–769.
48. Batey, R.T., Inada, M., Kujawinski, E., Puglisi, J.D. & Williamson, J.R. (1992). Preparation of isotopically labelled ribonucleotides for multidimensional NMR spectroscopy of RNA. *Nucleic Acids Res.* **20**, 4515–4523.
49. Batey, R.T., Battiste, J.L. & Williamson, J.R. (1995). Preparation of isotopically enriched RNAs for heteronuclear NMR. *Methods Enzymol.* **261**, 300–322.
50. Sklenár, V. & Bax, A. (1987). Spin-echo water suppression for the generation of pure phase two-dimensional NMR spectra. *J. Magn. Reson.* **74**, 469–479.
51. Parkinson, G., Voitechovsky, L., Clowney, L., Brünger, A.T. & Berman, H. (1996). New parameters for the refinement of nucleic acid-containing structures. *Acta Cryst. D* **52**, 57–64.